100.138

Spatial Searching for Solar Physics Data

J.A. Hourclé
NASA-GSFC (Wyle)
joseph.a.hourcle@nasa.gov

J. Spencer
NASA-GSFC (Telophase)
jennifer.spencer@nasa.gov

Virtual Solar Observatory
http://virtualsolar.org/

The Virtual Solar Observatory allows searching across many collections of solar physics data, but does not yet allow a researcher to search based on the location and extent of the observation, other than by selecting general categories such as full disk or off limb. High resolution instruments that observe only a portion of the the solar disk require greater specificity than is currently available.

We believe that finer-grained spatial searching will allow for improved access to data from existing instruments such as ***TRACE***, ***XRT*** and ***SOT***, and well as from upcoming missions such as ***ATST*** and ***IRIS***. Our proposed solution should also help scientists to search on the field of view of full-disk images that are out of the Sun-Earth line, such as STEREO/EUVI and observations from the upcoming ***Solar Orbiter*** and ***Solar Probe Plus*** missions.

We present our current work on cataloging sub field images for spatial searching so that researchers can more easily search for observations of a given feature of interest, with the intent of soliciting information about researcher's requirements and recommendations for further improvements.

Use Cases

Although there are other reasons for being interested in spatial coordinates, the VSO's goal is to index the data to improve searching. We are attempting to meet the following uses:

1. **Find what instruments were pointed at a given location**
   - 1a. **... stationary in helioprojective**
   - 1b. **... subject to rotation**
   - 1c. **... off the disk**
   - 1d. **... in heliographic with a given viewing angle**
2. **Find when two or more given instruments were observing the same region**
   - 2a. **... simultaneously**
   - 2b. **... from different viewing angles**

We assume that existing VSO search parameters, such as time and spectral ranges would be used in conjunction with spatial searches.

Example : What active regions were observed simultaneously by both EUVI and TRACE?

Helioprojective Cartesian

There are several different ways to track the spatial extent of an observation. Most FITS files will describe their observation in terms of the pointing (`CRPIXn` or `XCEN` / `YCEN`), plus the number of pixels (`NAXISn`), the size of each pixel (`CDELTn`), and the rotation (`CROTAn`).

Some data systems, such as HEK, allow you to search for `XCEN`/`YCEN` within a given range. Unfortunately, without the full bounding box, we can't identify if our region of interest might have been caught in an observation even if it wasn't necessarily the original observer's target.

To determine if there was any overlap, we must instead compute the bounding box. In the ideal case with no rotation:

```
x = XCEN ± ( NAXIS1 * CDELT1 )
y = YCEN ± ( NAXIS2 * CDELT2 )
```

HEK allows you to also search on the bounding box's lower left and upper right corners (`X1`/`Y1`, `X2`/`Y2`).

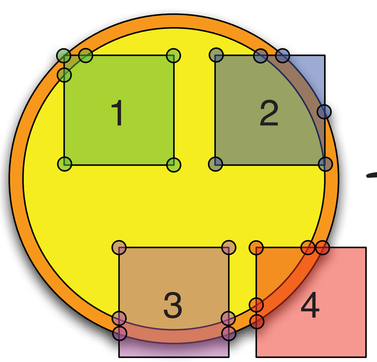
Unfortunately, a bounding box only in terms of SolarX/Y will have error at the edges for images that aren't aligned with the axis of rotation. We propose to instead calculate each corner and load them as polygons into one of the spatial databases commonly used by the GIS community.

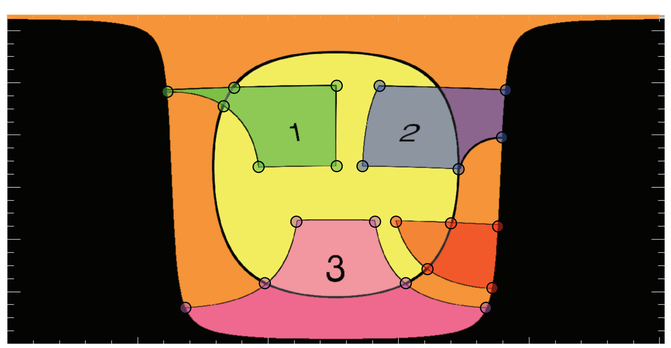
Helioprojective coordinates also have problems when searching for long-lived features on the solar disk; when looking for incidential observations that may happen at the edge of the field of view, we need to break up our query into multiple shorter segments so we can move our region of interest for each time step as we follow the feature.

Proposed Solution

To improve searching for long-lived features, we plan to index the observations in heliographic coordinates. Carrington reduces the movement of most features in the reference system and ***allows for tracking observations from not only away from the Sun-earth line, but significantly out of the plane of the ecliptic***.

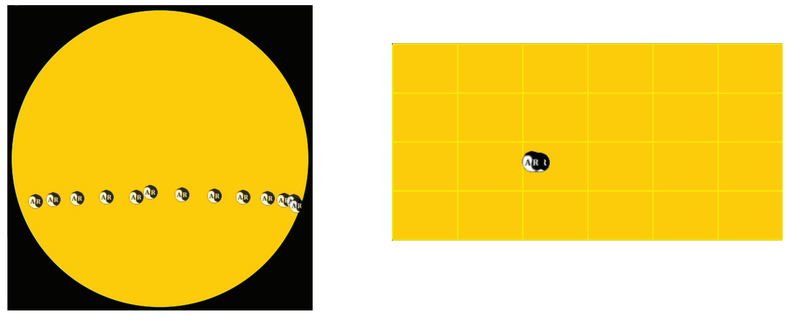
We need to test how to automate the processing to enable fast searching while minimizing the error from describing the observations as polygons.



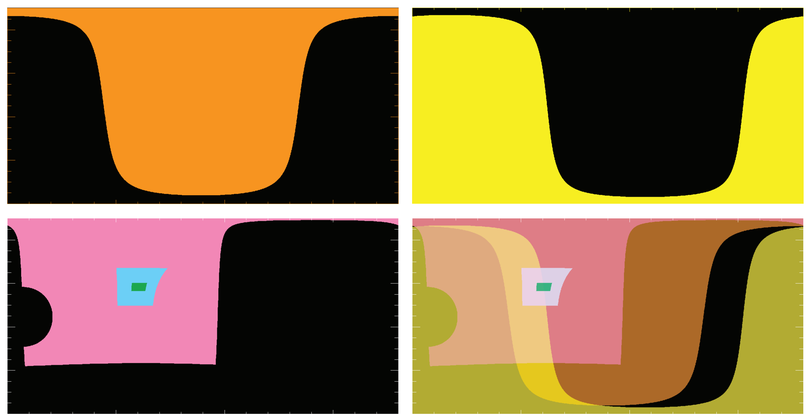


Limb Masked and Corners Identified in Helioprojective, then Reprojected to Heliographic Coordinates

Tracking AR#11743 in Helioprojective vs. Carrington Heliographic (via iSolSearch):



Tracking overlap near 2010-06-21 23:00 UT of EUVI (Ahead, top left; Behind, top right), SOT, TRACE and XRT (lower left, from center outward) in Carrington.



Challenges:

Reprojecting

When we transform an image to Carrington or Stonyhurst coordinates, they are no longer rectangular; attempting to describe the reprojection in terms of a bounding box introduces significant error.

Reprojecting at the Limb

Observations of the limb are difficult to project to other coordinate systems, especially for wavelengths in which the limb is not optically thick.

Off-Limb

Off-limb observations cannot easily be compared to observations from other locations without knowing their location or making a plane-of-sky assumption.

Temporal Proximity

When computing if two instruments viewed an event, we may need to have an indication of how close in time the observations must be.

Observations vs. Campaigns

With high cadence observations, it may not be efficient to track each observation individually. We may need to track fewer 'key frames', or use time as a third dimension and track whole campaigns as a volume.

Campaigns or other aggregations may be required to perform qualifiers on use case #2 such as 'with at least 30 images' or 'with a cadence better than 60 seconds' or 'overlapping for at least 15 minutes'

Questions still remaining:

Are we missing any spatial searching use cases?

Do we need to track observations in other coordinate systems?

How do we minimize false negatives when searching while still returning results in a timely manner and minimizing false positives?

Where is the limb? Can the community agree on a fixed distance for the limb, or is that something that we need to compute for each search?

If you're working on spatial searching, and would like to collaborate, or have experience with PostGIS or other spatial database, we'd love to talk to you.

References:

'FITS Keywords used by ZTOOLS v1.02', ftp://sohoftp.nascom.nasa.gov/solarsoft/packages/ztools/fits-definitions.html

'Heliophysics Event Registry web API', http://vso.stanford.edu/hekwiki/ApplicationProgrammingInterface

'iSolSearch', http://www.lmsal.com/isolsearch

Thompson, (2006). 'Coordinate Systems for Solar Image Data', http://adsabs.harvard.edu/abs/2006A%26A...449..791T